Deuxième épreuve CAPESA externe 1999

François Sauvageot

2 mars 2001

Question 0.1 On vérifie immédiatement que

$$(U|\lambda_1 V_1 + \lambda_2 V_2)_n = \lambda_1 (U|V_1)_n + \lambda_2 (U|V_2)_n$$

pour tout couple de scalaires (λ_1,λ_2) et tout triplet de vecteurs (U,V_1,V_2). On a aussi

$$(V|U)_n = \overline{(U|V)_n} \ .$$

La formule, pour U le vecteur colonne $(u_i)_{1\leq i\leq n}$,

$$U^*U = \sum_i |u_i|^2$$

montre que la forme sesqui-linéaire $(\ |\ )_n$ est bien définie positive et donc définit un produit hermitien sur $M_{n,1} \simeq \mathbf{C}^n$.

Question 0.2 Si U est le vecteur de colonne (x_i) et V le vecteur colonne $(\bar{y_i})$, l'inégalité demandée est juste l'inégalité de Cauchy-Schwartz.

Question 0.3 La seconde assertion résulte de la première puisque les deux quantités étudiées sont les complexes conjuguées de celles de la première assertion.

Pour la première, l'égalité pour $U = V - W$ donne $||V-W||_n^2 = 0$ et donc $V = W$.

Question 0.4 On écrit

$$(AU|V)_n = (AU)^*V = (U^*A^*)V = U^*(A^*V) = (U|A^*V)_n \ .$$

Question 0.5 Si U appartient au noyau de A et $V = A^*W$ à l'image de A^*, alors

$$(U|V)_n = (U|A^*W)_n = (AU|W)_n = (0|W)_n = 0 \ .$$

La somme entre $Ker(A)$ et $Im(A^*)$ est donc orthogonale et, *a fortiori*, directe. La dimension de cette somme est donc la somme des dimensions des deux sous-epaces et donc, par le théorème du rang, égale à la dimension de la source. Cette somme est donc tout $M_{p,1}$.

La seconde assertion résulte de la première appliquée à A^*.

Question 0.6.1

$$tr(AB) = \sum_{i,j} a_{ij}b_{ji} = \sum_{j,i} b_{ji}a_{ij} = tr(BA) \ .$$

Aussi, si $B = P^{-1}AP$ est semblable à A,

$$tr(B) = tr(P^{-1}(AP)) = tr((AP)P^{-1}) = tr(A) \ .$$

Question 0.6.2 Comme $\mathbf{C}$ est algébriquement clos, A est triangularisable. Il existe donc P inversible tel que $P^{-1}AP$ est triangulaire et admet les valeurs propres de A sur sa diagonale (car ces deux matrices ont même polynôme caractéristique : $\det(P^{-1}AP - XId) = \det(P^{-1})\det(A-XId)\det(P) = \det(A-XId)$). Il en résulte que la trace de A est celle de $P^{-1}AP$ et est donc la somme des valeurs propres de A.

Partie I

Question I.1.1 Ces deux matrices sont hermitiennes et donc diagonalisables.

Question I.1.2 Soit X un vecteur propre non nul pour AB pour la valeur propre ρ. Comme ABX n'est pas nul, BX ne l'est pas non plus. Or

$$BA(BX) = B(ABX) = \rho BX$$

et donc BX est vecteur propre pour BA, associé à la valeur propre ρ. On obtient la réciproque en échangeant les rôles de A et B, et donc AB et BA ont les mêmes valeurs propres non nulles.

Question I.1.3 Si $A^*AX = \lambda X$, alors

$$||AX||_n^2 = X^*(A^*AX) = \lambda ||X||_n^2$$

et donc λ est un réel positif (puisque $||X||_n^2$ est non nul).

En appliquant ce résultat à A^*, on obtient l'assertion pour AA^*.

Question I.1.4 Soit P hermitienne telle que

$$P^{-1}A^*AP = D$$

soit une matrice diagonale à coefficients positifs (les valeurs propres de A^*A). Soit D' la matrice diagonale dont les coefficients sont les racines carrées de ceux de D, la matrice $B = PD'P^{-1}$ répond à la question.

On obtient C en appliquant ce qui précède à A^*.

Question I.2 On a

$$A^* = \begin{pmatrix} 1 & 1 \\ -1 & -i \\ 0 & 0 \end{pmatrix}$$

et donc

$$A^*A = \begin{pmatrix} 2 & i-1 & 0 \\ -1-i & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad AA^* = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix}.$$

Le polynôme caractéritique de ces deux matrices est, respectivement,

$$P_{A^*A}(X) = -X(X^2 - 4X + 2) \qquad \text{et} \qquad P_{AA^*}(X) = X^2 - 4X + 2\,.$$

Les nombres singuliers de A sont donc $2+\sqrt{2}$ et $2-\sqrt{2}$.

Pour A^*A les vecteurs propres pour 0, $2+\sqrt{2}$ et $2-\sqrt{2}$ sont donnés par $x = y = 0$, $z = -\sqrt{2}x + (i-1)y = 0$ et $z = \sqrt{2}x + (i-1)y = 0$. Une base orthonormée de vecteurs propres est donc

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{i-1}{2} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{i-1}{2} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

Pour AA^* les vecteurs propres pour $2+\sqrt{2}$ et $2-\sqrt{2}$ sont donnés par $-\sqrt{2}x + (1+i)y = 0$ et $\sqrt{2}x + (1+i)y = 0$. Une base orthonormée de vecteurs propres est donc

$$\begin{pmatrix} \frac{i+1}{2} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{i+1}{2} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

Question I.3 On a

$$(A^{-1})^*A^{-1} = (AA^*)^{-1} \qquad \text{et} \qquad A^{-1}(A^{-1})^* = (A^*A)^{-1}$$

et donc les valeurs propres communes à $(A^{-1})^*A^{-1}$ et $A^{-1}(A^{-1})^*$ sont les valeurs propres communes aux inverses de A^*A et AA^* et sont donc les inverses des valeurs propres communes à ces deux dernières matrices. Ce sont donc les inverses des nombres singuliers de A.

Question I.4.1 Ces deux matrices sont adjointes, elles ont donc mêmes valeurs propres.

Question I.4.2 On a

$$\sum_k \mu_k^2 = tr(A^*A) \qquad \text{et} \qquad \prod_k \mu_k = \sqrt{\det(A^*A)} = \sqrt{\det(A)\overline{\det(A)}} = |\det(A)| \ .$$

Comme le produit des valeurs propres de A est égal à son déterminant, on a immédiatement

$$\prod_k |\lambda_k| = |\det(A)| = \prod_k \mu_k \ .$$

Pour la seconde assertion, observons que l'on peut toujours trianguliariser A et même, en orthonormalisant la base de triangularisation par le procédé de Gram-Schmidt, on peut trianguliariser dans une base orthonormale. En termes matriciels, cela veut dire qu'il existe une matrice unitaire U telle que U^*AU est une matrice triangulaire T. On a donc $A = UTU^* = UTU^{-1}$. Les valeurs propres de A sont aussi celles de T, i.e. les éléments diagonaux de T. Donc $\sum_i |\lambda_i|^2$ est la somme des carrés des modules éléments diagonaux de T.

De plus $\sum_i \mu_i^2$ est la somme des valeurs propres de A^*A, donc sa trace. On calcule

$$tr(A^*A) = \sum_{i,j} \bar{a_{ji}} a_{ji}$$

et donc la trace de A^*A est la somme des carrés des modules de tous les termes de la matrice. Remarquons que

$$tr(A^*A) = tr(UT^*U^*UTU^*) = tr(UT^*TU) = tr(UT^*TU^{-1}) = tr(T^*T)$$

et donc $\sum_i \mu_i^2$ est la somme des carrés des modules de tous les éléments de T. Cette quantité est donc supérieure à la somme des carrés des modules des éléments diagonaux et donc

$$\sum_i |\lambda_i|^2 \leq \sum_i \mu_i^2 \ .$$

Question I.5 Remarquons que, pour tout vecteur colonne X, on a

$$AX = 0 \Rightarrow A^*AX = 0$$

et

$$A^*AX = 0 \Rightarrow X^*A^*AX = 0 \Rightarrow ||AX||_n^2 = 0$$

et donc $Ker(A) = Ker(A^*A)$. De même $Ker(A^*) = Ker(AA^*)$.

On prend pour X_k une base orthonormée de vecteurs propres de A^*A. On les ordonne de sorte que X_k corresponde à $\lambda_k = \mu_k^2$. D'après la remarque précédente, les X_k pour $k > q$ forment un base du noyau de A^*A et donc du noyau de A. Il en résulte que les AX_k, pour $k \leq q$ forment une base de l'image de A.

On prend $Y_k = AX_k/\mu_k$ pour $k \leq q$. Ces Y_k forment donc une base de l'image de A. Un supplémentaire de ce sous-espace est donc donné par le noyau de A^*. Prenons donc, pour $k > q$, des Y_k formant une base orthonormée de $Ker(A^*)$.

Ces vecteurs vérifient toutes les propriétés requises puisque

$$A^*Y_k = A^*AX_k/\mu_k = \lambda_k X_k/\mu_k = \mu_k X_k$$

et

$$||Y_k||_n^2 = (AX_k|AX_k)_n/\mu_k^2 = (X_k|A^*AX_k)_n/\lambda_k = ||X_k||_n^2 = 1 \ .$$

Question I.6 Soit Q la matrice envoyant la base canonique sur les X_i et P celle qui envoie les Y_i sur la base canonique. Ces deux matrices existent et sont unitaires puisque les X_i et les Y_i forment des bases orthonormales.

On a $PAQe_i = PAX_i = \mu_i PY_i = \mu_i e_i$, si (e_i) est la base canonique de $M_{n,1}$. Donc PAQ a bien les coefficients voulus.

Partie II

Question II.1 C'est une reformulation de $Im(A) \oplus^{\perp} Ker(A^*) = M_{n,1}$.

Question II.2.1 On applique la seconde égalité de 0.5 à A^*A et donc

$$M_{p,1} = Ker(AA^*) \oplus Im(A^*A) \ .$$

En conséquence, comme $Ker(AA^*) = Ker(A^*)$, $Im(A^*A) \subset Im(A^*)$ et

$$rg(A^*) = p - dim(Ker(A^*))$$

on doit avoir

$$Im(A^*A) = Im(A^*) \ ,$$

ce qui est l'assertion recherchée.

Question II.2.2 Comme A^*A est diagonalisable, ses sous-espaces caractéristiques sont ses sous-espaces propres. En particulier $Ker(A^k) = Ker(A)$ pour tout $k > 0$. Il en résulte donc $Im(A^k) = Im(A)$ et aussi $Im(A^k) = Im(A^j)$ pour tout couple d'entiers (j,k) strictement positifs. Ce qui est bien l'assertion demandée.

Question II.2.3.1 Si X est solution de (2), alors $\Phi_Y(X) = 0$ et, comme Φ_Y est positive, c'est bien un de ses minima.

Question II.2.3.2 Comme $(\ |\ )_n$ est un produit hermitien défini positif et que AX décrit $Im(A)$ quand X décrit $M_{p,1}$, le minimum de Φ_Y est atteint quand AX est la projection orthognale de Y sur $Im(A)$. D'après 0.5 ceci est équivalent à ce que $Y - AX$ appartienne à $Ker(A^*)$, ce qui est exactement l'assertion.

Question II.2.3.3 On sait que l'équation précédente admet une solution d'après II.2.1. Soit X_0 une telle solution. Si X en est une autre, on a $AX = AX_0$ est la projection orthogonale de Y sur $Im(A)$. Donc $X - X_0$ appartient à $Ker(A)$. Il existe un unique X appartenant à $Im(A^*)$ (l'orthogonal de $Ker(A)$) et tel que $X - X_0$ soit dans $Ker(A)$, c'est la projection orthogonale de X_0 sur $Im(A^*)$. Cette projection est unique et de plus ne dépend pas du X_0 choisi. D'où l'assertion.

Question II.2.3.4 ϕ_A est linéaire puisque l'équation en X, $A^*AX = AY$ est linéaire en Y et l'orthogonalité à $Ker(A)$ est aussi linéaire :

$$A^*AX = Y \qquad A^*AX' = Y' \Rightarrow A^*A(aX + a'X') = aY + a'Y'$$

et

$$X,X' \in Im(A^*) \Rightarrow aX + a'X' \in Im(A^*) \ ,$$

d'où

$$\phi_A(aY + a'Y') = aX + aX' = a\phi_A(Y) + a'\phi_A(Y') \ .$$

On a

$$\phi_A(Y) = 0 \Leftrightarrow A^*Y = 0$$

et donc $Ker(\phi_A) = Ker(A^*)$. On a $Im(\phi_A) \subset Im(A^*)$ par construction de ϕ_A, l'égalité résulte de l'égalité des noyaux, donc des dimensions de ces sous-espaces.

Question II.2.4.1 Les pseudo-solutions sont définies modulo $Ker(A)$, il y a donc unicité si et seulement si A est injective, ou encore si A est de rang 2. C'est le cas puisque les x_i ne sont pas tous égaux.

Question II.2.4.2 Comme $Y - AX_0$ admet pour coordonnées les $y_i - a_0x_i - b_0$, le vecteur X_0 est tel que

$$\sum_i |y_i - a_0x_i - b_0|^2$$

est minimal. La droite $y = ax_0 + b_0$ est donc la droite de la méthode des moindres carrés.

Remarques : il ne me semble pas qu'elle minimise la somme des carrés des distances à une droite, puisque le carré de la distance à $y = a_0x + b_0$ est donné par $(y - a_0x - b_0)^2/(a_0^2 + 1)$. De plus on n'a montré nulle part que a_0 et b_0 sont réels.

Partie III

Question III.1.1 Remarque: il y a évidemment une faute dans l'énoncé, la seconde inégalité étant une majoration de $||AU||_n$ et non de $||AU||$ (ce qui n'aurait aucun sens).

Montrons d'abord la seconde propriété. Pour U non nul et $u = ||U||_n$, posons $V = U/u$. C'est un vecteur de norme 1 et on a

$$||AU||_n = ||A(uV)||_n = u||AV||_n \leq u||A||_S = ||A||_S.||U||_n \ .$$

Pour U quelconque, on a donc

$$||ABU||_n \leq ||A||_S.||BU||_n \leq ||A||_S.||B||_S.||U||_n$$

et donc

$$||AB||_S \leq ||A||_S.||B||_S \ .$$

Question III.1.2 On a

$$||AU||_n^2 = U^*(A^*AU) \ .$$

Comme la matrice A^*A est hermitienne, on peut la diagonaliser dans une base orthonormée. Soit λ_i ses valeurs propres et U_i la base associée. On a $U = \sum_i u_i U_i$ et

$$||AU||_n^2 = \sum_i \lambda_i |u_i|^2 \leq \max(\lambda_i) \sum_i |u_i|^2 = \max(\lambda_i)||U||_n^2$$

avec égalité si et seulement si U est vecteur propre de A^*A pour la valeur propre $max(\lambda_i)$. Il en résulte que le carré de $||A||_S$ est cette valeur propre maximale de A^*A et donc le carré du nombre singulier maximal de A. D'où l'assertion.

Question III.2.1 Je présume qu'il faudrait calculer ... Voyons une autre méthode!

La formule pour la trace de AB donnée en 0.6.1 montre que

$$tr(A^*A) = ||A||_E^2 \ .$$

La première inégalité est donc équivalente à

$$tr(B^*A^*AB) \leq tr(A^*A)tr(B^*B)$$

ou encore

$$tr(BB^*A^*A) \leq tr(A^*A)tr(BB^*) \ .$$

Or, on a affaire à deux matrices hermitiennes. On peut donc diagonaliser la première dans une base orthonormale (U_i), chacun des U_i étant associé à la valeur propre λ_i. Comme on est dans une base orthonormale $U_i^*MU_i$ est exactement le coefficient d'indice (i,i) de la matrice M, pour toute matrice M exprimée dans la base (U_i). La trace de M est donc la somme sur i de ces termes. De plus si $M = BB^*$, alors ce terme vaut $||B^*U_i||_n^2$. On a donc

$$tr(BB^*AA) = \sum_i U_i^*BB^*A^*AU_i = \sum_i \lambda_i ||B^*U_i||_n^2 \ .$$

Comme tous ces termes sont positifs, cette somme est évidemment majorée par

$$\left(\sum_j \lambda_j\right) . \left(\sum_i ||B^*U_i||_n^2\right) = tr(A^*A)tr(BB^*) \ .$$

Il en résulte $||AB||_E \leq ||A||_E.||B||_E$.

En utilisant le calcul de $||AU||_n^2$ dans une base de diagonalisation de A^*A, que l'on vient de faire, on obtient

$$||AU||_n^2 = \sum_i \lambda_i |u_i|^2 \leq (\sum_j \lambda_j).(\sum_i |u_i|^2) = tr(A^*A)||U||_n^2 \ .$$

D'où la seconde inégalité.

On a $||I_n||_E = tr(I_n) = n$.

Question III.2.2 On a déjà montré $||A||_E = tr(A^*A)$ et la formule demandée résulte donc de l'égalité démontrée en I.4.2. On a donc

$$||A||_E^2 \geq \max(\mu_k^2) = \max(\mu_k)^2 = ||A||_S^2$$

et l'inégalité demandée en résulte.

Il faut noter que cette inégalité est immédiate au vu de la seconde propriété imposée aux normes : $||\ ||_S$ est la plus petite norme à vérifier cela. Comme elle vérifie la première inégalité aussi, elle est évidemment la plus petite norme à vérifier les deux. Ici « la plus petite » est à prendre au sens fort : elle est minimale pour tout A.

Partie IV

Question IV.1 Si on se sert de $||I_n|| \geq ||I_n||_S$, c'est immédiat. Mais on peut aussi dire que pour tout A, on a

$$||AI_n|| \leq ||I_n||.||A|| \ .$$

L'inégalité demandée en résulte en prenant $||A||$ non nulle, i.e. A non nulle.

Question IV.2 Il suffit pour cela de prendre U un vecteur non nul propre pour A et associé à la valeur propre λ, puis d'écrire la seconde inégalité :

$$||\lambda U||_n \leq ||A||.||U||_n \ .$$

Question IV.3.1 Une matrice carrée est inversible si et seulement si elle n'admet pas 0 pour valeur propre. Donc $I_n + A$ est inversible si et seulement si -1 n'est pas valeur propre de A. Cela résulte de $||A|| < 1$ et de IV.2.

Question IV.3.2 Par sommes télescopiques, on a

$$(I_n + A) \cdot \left(\sum_{k=0}^{p} (-1)^k A^k \right) = I_n + (-1)^p A^{p+1} \ .$$

(Ce n'est pas précisé, mais on convient évidemment que $A^0 = I_n$.)

Il va falloir évaluer la norme de $(I_n + A)^{-1}$. Si on ne veut pas être embêté par $||I_n||$ dans ce calcul, voilà comment faire : soit X une matrice quelconque, on a

$$X = (I_n + A)(I_n + A)^{-1}X = (I_n + A)^{-1}X + A(I_n + A)^{-1}X$$

et donc

$$(I_n + A)^{-1}X = X - A(I_n + A)^{-1}X \ .$$

Il en résulte (par l'inégalité triangulaire : n'oublions pas que $||\ ||$ est aussi une norme !)

$$||(I_n + A)^{-1}X|| \leq ||X|| + ||-A(I_n + A)^{-1}X|| \leq ||X|| + ||A||.||(I_n + A)^{-1}X||$$

et donc

$$||(I_n + A)^{-1}X|| \leq \frac{||X||}{1 - ||A||} \ .$$

On applique cela à la matrice X choisie ainsi

$$X = (I_n + A) \cdot \left(\sum_{k=0}^{p} (-1)^k A^k \right) - I_n = (-1)^p A^{p+1} \ .$$

En prenant la première expression de X, on tombe sur le membre de gauche de l'inégalité recherchée et en prenant la seconde expression, on tombe sur le terme de droite.

On en conclut que la série considérée converge normalement vers l'inverse de $I_n + A$ et donc converge tout court car M_n est complet pour toute norme.

Question IV.4 Si ϵ_A vérifie l'inégalité, alors $||A^{-1}\epsilon_A|| < 1$ et donc $I_n + A^{-1}\epsilon_A$ est inversible, il en est de même de son produit par n'importe quelle matrice inversible, par exemple A et donc $A + \epsilon_A$ est inversible.

On a donc

$$(A+\epsilon_A)^{-1} = A^{-1}(I_n + \epsilon_A A^{-1})^{-1} = A^{-1}\sum_{k=0}^{+\infty}(-1)^k(\epsilon_A A^{-1})^k$$

et

$$(A+\epsilon_A)^{-1} - A^{-1} = A^{-1}\sum_{k=1}^{+\infty}(-1)^k(\epsilon_A A^{-1})^k \ .$$

D'où

$$\begin{aligned} ||(A+\epsilon_A)^{-1} - A^{-1}|| &\leq ||A^{-1}||\sum_{k=1}^{+\infty}||A^{-1}||^k||\epsilon_A||^k \\ &\leq ||A^{-1}||\frac{||A^{-1}||.||\epsilon_A||}{1-||A^{-1}\epsilon_A||} \\ &\leq \frac{||A^{-2}||.||\epsilon_A||}{1-||A^{-1}||.||\epsilon_A||} \end{aligned}$$

Question IV.5 Quand A est inversible (ce qui est, ici, affirmé), la matrice $\det(A)A^{-1}$ est la transposée de la comatrice de A. Le calcul des mineurs est fastidieux mais facile et on trouve

$$\det(A)A^{-1} = \begin{pmatrix} -0.1263 & 0.2548 & -0.0908 \\ -0.2467 & 1.9102 & -0.6272 \\ 0.2806 & -1.9516 & 0.5706 \end{pmatrix} .$$

On calcule aussi $\det(A) = -0.11363$.

On va appliquer ce qui précède avec $A = A$, $\epsilon_A = B - A$ et $||\ || = ||\ ||_E$.

On a $||\epsilon_A||_E < \sqrt{9.10^{-4}} = 3.10^{-2}$. On aimerait donc que $||A^{-1}||_E$ soit inférieur à $100/3$. Ou encore que $||\det(A)A^{-1}||_E$ soit inférieur à $100|\det(A)|/3 = 11.363/3$. Il suffit donc de majorer par 3.787. On aimerait donc que la somme des carrés des coefficients de $\det(A)A^{-1}$ soit inférieure à 3.787^2, donc par exemple à 14. Or on trouve 8.40528742. La matrice B est donc inversible.